

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2004, 15:22:32; Search time 182 Seconds

(without alignments)
7998.001 Million cell updates/sec

Title: US-09-989-862-284

Perfect score: 2623

Sequence: 1 ttgagcgacgagtgagctctct.....taaaaaaaaaaaaaaaaaa 2623

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1113.6	42.5	1621	2	US-09-014-969-18
2	443.4	16.9	482	4	US-09-621-976-53
3	300.6	11.5	2330	4	US-09-800-729-12
4	108.4	4.1	3486	4	US-09-438-906-1
5	108.4	4.1	3486	4	US-09-438-906-3
6	54.2	2.1	1438	3	US-09-187-331-4
7	54.2	2.1	1438	4	US-09-470-946-4
8	54.2	2.1	3104	1	US-08-346-455B-66
9	54.2	2.1	3104	3	US-08-977-221-66
10	54.2	2.1	3104	5	PCT-US95-06613-66
11	54.2	2.1	3104	5	PCT-US95-06613-66
12	53.4	2.0	3001	4	US-09-539-333D-215
13	53.4	2.0	90541	4	US-09-759-359A-3
14	52.4	2.0	3000	1	US-08-393-985-3
15	51.2	2.0	26000	4	US-09-843-376-10
16	51.2	1.9	2712	1	US-08-346-455B-37
17	51.2	1.9	2712	3	US-08-977-221-37
18	51.2	1.9	2712	4	US-09-483-831B-37
19	51.2	1.9	2712	5	PCT-US95-06613-37
20	51.2	1.9	2946	3	US-08-346-455B-35
21	51.2	1.9	2946	3	US-08-977-221-35
22	51.2	1.9	2946	4	US-09-483-831B-35
23	51.2	1.9	2946	5	PCT-US95-06613-35
24	51.2	1.9	3251	1	US-08-346-455B-68
25	51.2	1.9	3251	4	US-08-977-221-68
26	51.2	1.9	3251	4	US-09-483-831B-68
27	51.2	1.9	3251	5	PCT-US95-06613-68

28	51	1.9	6350	2	US-08-385-335A-9	Sequence 9, Appl1
29	50.6	1.9	4771	2	US-08-866-650-2	Sequence 2, Appl1
30	50.6	1.9	4771	2	US-09-021-287-2	Sequence 2, Appl1
31	50.6	1.9	4771	3	US-09-240-473-2	Sequence 2, Appl1
32	50.4	1.9	264	1	US-08-222-177A-10	Sequence 10, Appl1
33	50	1.9	11049	4	US-10-204-708-23	Sequence 23, Appl1
34	49.4	1.9	2614	4	US-09-052-753B-8	Sequence 8, Appl1
35	49	1.9	3474	4	US-09-527-236A-1	Sequence 1, Appl1
36	49	1.9	3474	4	US-09-756-854-1	Sequence 1, Appl1
37	48.6	1.9	1016	1	US-08-109-391A-3	Sequence 3, Appl1
38	48.6	1.9	1016	2	US-08-459-012A-3	Sequence 3, Appl1
39	48.6	1.9	1016	3	US-08-459-012A-3	Sequence 3, Appl1
40	48.6	1.9	1016	3	US-08-458-860A-3	Sequence 3, Appl1
41	48.6	1.9	51259	3	US-08-781-891-209	Sequence 209, App
42	48.6	1.9	51259	4	US-09-618-166-209	Sequence 209, App
43	48.4	1.8	324	1	US-08-222-177A-36	Sequence 36, Appl1
44	48.4	1.8	1936	4	US-09-566-921-41	Sequence 41, Appl1
45	48.4	1.8	8387	2	US-08-532-814-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-014-969-18
Sequence 18, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallie, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Apostolou, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1621 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-18
Query Match 42.5%; Score 1113.6; DB 2; Length 1621;
Best Local Similarity 94.7%; Pred. No. 4.4e-298;
Matches 1188; Conservative 0; Mismatches 4; Indels 62; Gaps 1;

```
QY 1 TTGACGCGAGGTGAGCTCTGAGGCTTCGCGGGGGCTTCCTCCAGTCAACCTCCCGCGGT 60
DB 14 TTGACGCGAGGTGAGCTCTGAGGCTTCGCGGGGGCTTCCTCCAGTCAACCTCCCGCGGT 73
QY 61 TACCGCGCGCGCGCGAGGAGTCTCTCCAGACCTCCCTCCCGTGGCTCAAACTAA 120
DB 74 TACCGCGCGCGCGCGAGGAGTCTCTCCAGACCTCCCTCCCGTGGCTCAAACTAA 133
QY 121 TACGAGCTGACCGAGTCTGCGAGGCTGGAGAGAAATTAAGGGGAGAAAGACAGAG 180
DB 134 TACGAGCTGACCGAGTCTGCGAG-----158
QY 181 AAGGACCTGACATCCATAGCCAGATGATTATCTTACGATGATGATCACTCTTG 240
DB 159 -----GATTATCTTACGATGATGATCACTCTTG 191
QY 241 AAATGACTTCGAAATTTATCTTGGTCTCTGATCTGCTGAGCTGATCTTCAAC 300
DB 192 AAATGACTTCGAAATTTATCTTGGTCTCTGATCTGCTGAGCTGATCTTCAAC 251
QY 301 ACCCTTCTCTCACTGACGACGAGAAAGGTTCTAGTCTCTTGTGATGATTCGT 360
DB 252 ACCCTTCTCTCACTGACGACGAGAAAGGTTCTAGTCTCTTGTGATGATTCGT 311
QY 361 TGGGATTAATTAATTAAGTTCAGAGCCCATTTCTATTAATTAATTAATTAAGT 420
DB 312 TGGGATTAATTAATTAAGTTCAGAGCCCATTTCTATTAATTAATTAATTAAGT 371
QY 421 CACGTGAAGCAAGTTCATATGTTTATTAACAAACCTACCTACCTATATCTTG 480
DB 372 CACGTGAAGCAAGTTCATATGTTTATTAACAAACCTACCTACCTATATCTTG 431
QY 481 GTTACTGGCTCTTTCAGAGATCATGAGATGATGATGATGATGATGATGATGAT 540
DB 432 GTTACTGGCTCTTTCAGAGATCATGAGATGATGATGATGATGATGATGATGAT 491
QY 541 CGGAACAAATCTTCTCCCTGGATCAGATGATGATGATGATGATGATGATGATGAT 600
DB 492 CGGAACAAATCTTCTCCCTGGATCAGATGATGATGATGATGATGATGATGATGAT 551
QY 601 GCGACACCAATATGATCAAAACAGAGGCGAGCATATATGATGATGATGATGATGAT 660
DB 552 GCGACACCAATATGATCAAAACAGAGGCGAGCATATATGATGATGATGATGATGAT 611
QY 661 CCGGAGACATGTAATAATACATPAAGCGCTTCTACTCATATCATGCTTACATGAG 720
DB 612 CCGGAGACATGTAATAATACATPAAGCGCTTCTACTCATATCATGCTTACATGAG 671
QY 721 TCAATTCATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 672 TCAATTCATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
QY 781 AATAATCTTGTCTTCTATTTGGAGAACCTGATGATGATGATGATGATGATGATGAT 840
DB 732 AATAATCTTGTCTTCTATTTGGAGAACCTGATGATGATGATGATGATGATGATGAT 791
QY 841 GACGCTCGCTCATGAGGCGCTGCTCATTTGATGATGATGATGATGATGATGATGAT 900
DB 792 GACGCTCGCTCATGAGGCGCTGCTCATTTGATGATGATGATGATGATGATGATGAT 851
QY 901 CAATGCTGAAAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 852 CAATGCTGAAAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
QY 961 GGAATGAGAGGAGTCTGAGGAGAGGTTAATGAACTTGAACAGTACCTGATGATGATGAT 1020
DB 912 GGAATGAGAGGAGTCTGAGGAGAGGTTAATGAACTTGAACAGTACCTGATGATGATGAT 971
QY 1021 CACTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 972 CACTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
```

```
QY 1081 GATGAGCTATGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1032 GATGAGCTATGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
QY 1141 GAGCTTCAAGAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1092 GAGCTTCAAGAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
QY 1201 GCTGATGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
DB 1152 GCTGATGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205

RESULT 2
US-09-621-976-53
; Sequence 53, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Ubert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.05492
; CURRENT APPLICATION NUMBER: US/09/621.976
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 53
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 74..481
; NAME/KEY: sig peptide
; LOCATION: 74..139
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 9
; OTHER INFORMATION: seq FILAUSLSTPS/LQ
US-09-621-976-53

Query Match 16.9%; Score 443.4; DB 4; Length 482;
Best Local Similarity 99.8%; Pred. No. 7e-113;
Matches 444; Conservative 0; Mismatches 1; Indels 0;
Gaps 0;

QY 208 GATTATCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
DB 38 GATTATCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97
QY 268 TCCCTCATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
DB 98 TCCCTCATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 157
QY 328 AAGGTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
DB 158 AAGGTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 217
QY 388 CCCCATTTCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
DB 218 CCCCATTTCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 277
QY 448 ATTAACAAACCTACCTTACATTAATGATGATGATGATGATGATGATGATGATGATGAT 507
DB 278 ATTAACAAACCTACCTTACATTAATGATGATGATGATGATGATGATGATGATGATGAT 337
QY 508 GGGATTTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
DB 338 GGGATTTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 397
QY 568 AAGATATTTATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
DB 398 AAGATATTTATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 457
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2004, 20:34:08 ; Search time 105 Seconds

(without alignment)
2521.063 Million cell updates/sec

Title: US-09-989-862-285

Perfect score: 2561
Sequence: 1 WTSKFLVSTLALSLST.....SGLPALQMAHIAQPLLA 477

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-G/cgn2 1/USPTO.spool/US09989862.rnmat 19042004 120333 8715/app.query.fasta_1.647
-DB=Issued Patents NA -QEXT=fastap -SUFFIX=crni -NIMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=nnnn40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09989862 @CGN 1.1 56 @rnat 19042004 120333 8715 -NCP=6 -ICPU=3
-NO MAP -LARGEJITTER -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2 6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2 6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2 6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2 6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2 6/ptodata/2/ina/PTCUS.COMB.seq:*
6: /cgn2 6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1810	70.7	1621	2	US-09-014-969-18
2	1034	40.4	2330	4	US-09-800-729-12
3	719	28.1	482	4	US-09-621-976-53
4	629.5	24.6	3486	4	US-09-438-906-1
5	629.5	24.6	3486	4	US-09-438-906-3
6	534.5	20.9	1438	3	US-09-187-331-4
7	534.5	20.9	1438	4	US-09-470-946-4
8	515.5	18.1	2586	4	US-09-221-017B-448
9	474.5	18.5	2946	1	US-08-346-455B-35
10	474.5	18.5	2946	3	US-08-977-221-35
11	474.5	18.5	2946	4	US-09-483-831B-35
12	474.5	18.5	2946	5	PCT-US95-06613-35

13	474.5	18.5	3251	1	US-08-346-455B-68	Sequence 68, Appl
14	474.5	18.5	3251	3	US-08-977-221-68	Sequence 68, Appl
15	474.5	18.5	3251	4	US-09-483-831B-68	Sequence 68, Appl
16	474.5	18.5	3251	5	PCT-US95-06613-68	Sequence 68, Appl
17	454	17.7	3104	1	US-08-346-455B-66	Sequence 66, Appl
18	454	17.7	3104	3	US-08-977-221-66	Sequence 66, Appl
19	454	17.7	3104	4	US-09-483-831B-66	Sequence 66, Appl
20	454	17.7	3104	5	PCT-US95-06613-66	Sequence 66, Appl
21	366.5	14.3	2712	1	US-08-346-455B-37	Sequence 37, Appl
22	366.5	14.3	2712	3	US-08-977-221-37	Sequence 37, Appl
23	366.5	14.3	2712	4	US-09-483-831B-37	Sequence 37, Appl
24	366.5	14.3	2712	5	PCT-US95-06613-37	Sequence 37, Appl
25	197	7.7	506	4	US-09-621-976-31	Sequence 31, Appl
26	181.5	7.1	486	4	US-09-621-976-3282	Sequence 3282, Ap
27	111.5	4.4	1770	4	US-09-543-681A-22	Sequence 22, Appl
28	107.5	4.2	3111	4	US-09-421-017B-922	Sequence 11, Appl
29	105.5	4.1	2531	4	US-09-468-656A-11	Sequence 11, Appl
30	103.5	4.0	1749	1	US-07-649-591B-2	Sequence 2, Appl
31	103.5	4.0	1749	1	US-08-277-540-2	Sequence 2, Appl
32	103.5	4.0	1749	1	US-08-430-787A-2	Sequence 2, Appl
33	103.5	4.0	2472	3	US-08-335-844A-7	Sequence 7, Appl
34	103.5	4.0	2472	4	US-09-129-366-7	Sequence 7, Appl
35	103.5	4.0	3358	3	US-08-335-844A-20	Sequence 20, Appl
36	103.5	4.0	3358	4	US-09-129-366-20	Sequence 20, Appl
37	102.5	4.0	1480	4	US-09-124-864-11	Sequence 11, Appl
38	102.5	4.0	1480	4	US-08-961-527-243	Sequence 23, Appl
39	102	4.0	1359	4	US-07-621-670-1	Sequence 1, Appl
40	101	3.9	1625	4	US-09-813-133A-1	Sequence 1, Appl
41	101	3.9	1884	4	US-09-328-352-3421	Sequence 3421, Ap
42	99.5	3.9	2750	4	US-09-576-594-376	Sequence 376, Appl
43	99	3.9	1830121	4	US-09-557-884-1	Sequence 1, Appl
44	99	3.9	1830121	4	US-09-643-990A-1	Sequence 1, Appl
45	98.5	3.8	2290	3	US-08-961-083-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1
US-09-014-969-18
Sequence 18, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: McCoy, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Metberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1621 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-18

Alignment Scores:

Pred. No.:	8,71e-211	Length:	1621
Score:	1810.00	Matches:	335
Percent Similarity:	99.41%	Conservative:	0
Best Local Similarity:	99.41%	Mismatches:	2
Query Match:	70.68%	Indels:	0
DB:	2	Gaps:	0

US-09-989-862-285 (1-477) x US-09-014-969-18 (1-1621)

QY 1 MetThrSerIysPheIleuValSerPheIleuValalaIeuSerLeuSerThrThr 20
DB 195 ATGACTTCGAAATTTATCTGGTGTCTTCATCTGTCGACAGATGCTTTCAACAC 254
QY 21 PheSerLeuGlnLeuAspGlnGlnValLeuLeuValSerPheAspGlyPheArgTTP 40
DB 255 TTTTCTCCCAACCAACACAGCAAGGTTCTAGTTCCTTTTGATGATTCGTTGG 314
QY 41 AspTyrLeuTyrIysValIProThrProHisPheHisTyrIleMetIysTyrGlyValHis 60
DB 315 GATTAATTAATAAGTTCACAGCCCATTTCAATTAATAATAATAATAATAATAAGTTCAC 374
QY 61 ValIysGlnValThrAsnValPheIleThrIysThrTyrProAsnHisTyrThrIleuVal 80
DB 375 GTGAAGCAAGTACTATCTTTTATTAACAAAACCTACCTAACCTATTAATCTTTGTA 434
QY 81 ThrGlyLeuPheAlaGlnAsnHisGlyTyleValAlaAsnAspMetPheAspProIleArg 100
DB 435 ACTGACCTCTTTCAGAGATCATGAGTGTTCGAAAGATATGTTGATCTCATATTCG 494
QY 101 AsnIysSerPheSerLeuAspHisMetAsnHisTyrAspSerIysPheTropGlnGlnAla 120
DB 495 AACAAATCTTCTCTGATCAACATGAATTTTATGATTCAGTTTGGAGAGAGCG 554
QY 121 ThrProIleTPIleThrAsnGlnArgAlaGlyHisThrSerGlyAlaAlaMetTropPro 140
DB 555 AACCAATATGATGATCAACAAACAGAGGCGAGCAATACATGATGATGATGATGATG 614
QY 141 GlyThrAspValIysIleHisIysArgPheProThrHisTyrMetProTyrAsnGlnIser 160
DB 615 GGAACAGATGTAATAAATACATAGCGCTTTCATCATTAATGATGATGATGATGATG 674
QY 161 ValSerPheGlnAspArgValAlaIysIleValGlnTropPheThrSerIysGlnIser 180
DB 675 GTTCAATTTGAAGTATGAAGTTCGAAATGTTGATGATGATGATGATGATGATGATG 734
QY 181 AsnLeuGlyLeuLeuTyrTropGlnAspProAspMetGlyHisIleHisLeuGlyProAsp 200
DB 735 AATCTTGTCTTCTCTATTTGGGAAACCTGATGATGATGATGATGATGATGATGATG 794
QY 201 SerProLeuMetGlyProValIleSerAspIleAspIysLeuGlyTyrLeuIleGln 220
DB 795 AGTCGCTCATGAGGCTCTGCAATTCATGATTTGACAGAGATTTGATGATTTGATG 854
QY 221 MetLeuIysIysAlaIysLeuTyrAsnThrIleuAsnLeuIleIleThrSerAspHisGly 240
DB 855 ATGCGAAGAAAGCGAAAGTGTGGAACCTGTAACCTATATATCAAGATGATGATG 914
QY 241 MetThrGlnIysSerGlnGlnArgLeuIleGlnLeuAspGlnIleTyrIleuAspIysAspHis 260
DB 915 ATAGCGAGTCTCTGAGAAAGGTTAATAAATTTGACCAATTTGATGATGATGATGATG 974

QY 261 TyrThrIleuIleAspGlnSerProValAlaIleIleuProIysGlnGlyIysPheAsp 280
DB 975 TATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1034
QY 281 GluValIyrGlnAlaLeuThrHisAlaHisProAsnLeuThrValTyrIysGlnIser 300
DB 1035 GAAGTTATGAAGCATTAATCTGAGCTCATCTTAATCTTAATCTTAATCTTAATCTTA 1094
QY 301 ValProGlnArgTyrHisTyrIysTyrAsnSerArgIleGlnProIleIleAlaValAla 320
DB 1095 GTTCAGAAAGTGGCATTAATAAATACACATGCAATTCATCAATCAATCAATCAATG 1154
QY 321 AspGlnGlyTropHisIleLeuGlnAsnIysSerAspPheIleuGly 337
DB 1155 GATGAAGGATGGCAATTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1205

RESULT 2

US-09-800-729-12

Sequence 12, Application US/09800729

Patent No. 6605592

GENERAL INFORMATION:

APPLICANT: NI et al.

TITLE OF INVENTION: 32 Human secreted proteins

FILE REFERENCE: P2044P1

CURRENT APPLICATION NUMBER: US/09/800,729

PRIOR FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: PCT/US00/26013

PRIOR FILING DATE: 2000-09-22

PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 217

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 12

LENGTH: 2330

TYPE: DNA

ORGANISM: Homo sapiens

US-09-800-729-12

Alignment Scores:

Pred. No.:	9.81e-116	Length:	2330
Score:	1034.00	Matches:	208
Percent Similarity:	62.03%	Conservative:	73
Best Local Similarity:	45.92%	Mismatches:	132
Query Match:	40.37%	Indels:	40
DB:	4	Gaps:	10

US-09-989-862-285 (1-477) x US-09-800-729-12 (1-2330)

QY 6 IleLeuValSerPheIleuValalaIeuSerLeuSerThrThrPheSerLeuGlnIleu 25
DB 265 CTTATATTTTTCGATGATTAATCTGTTGTGAGGTACTCTTCCATAGCTG----- 318
QY 26 AspGlnGlnValIleuLeuValSerPheAspGlyPheArgTyrAspTyrLeuTyrIys 45
DB 319 ---CATCAAGCTGCTGCTGCTGCTTTCATGATGATGATGATGATGATGATGATG 375
QY 46 ValProThrProHisPheHisTyrIleMetIysTyrGlyValHisValIysGlnValThr 65
DB 376 TATGAATTTCCATCTCCAGATTTTATCAAGAGAGGCTCTGATGAGCATGATTA 435
QY 66 AsnValPheIleThrIysThrTyrProAsnHisTyrThrIleuValThrGlyLeuPheAla 85
DB 436 AATGTTTTTATCAAAAACATTTCTTAACATCACTACAGATGATGATGATGATGATG 495
QY 86 GluAsnHisGlyIleValAlaAsnAspMetPheAspProIleArgAsnIysSerPheSer 105
DB 496 GAAAGTCATGAGCATGCTGCTATTTCCATGATGATGATGATGATGATGATGATG 546
QY 106 LeuAspHisMetAsnIleTyrAspSerIys-----PheTyrGlnGlnAlaThrPro 122
DB 547 -----CATTTTTCGATTTGATGATGATGATGATGATGATGATGATGATGATG 600
QY 123 IleThrIleThrAsnGln---ArgAlaGlyHisThrSerGlyAlaAlaMetTropProGly 141

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 14:25:10 ; Search time 23 Seconds
(without alignments)
1070.678 Million cell updates/sec

Title: US-09-989-862-285

Perfect score: 2561
Sequence: 1 MTSKILVSTLAALSLSTT.....SQPLALQMAHAIQPLA 477

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCUTS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1810	70.7	355	2	US-09-014-969-19
2	1034	40.4	453	4	US-09-800-729-83
3	719	28.1	136	4	US-09-621-976-3913
4	629.5	24.6	873	3	US-09-187-331-6
5	629.5	24.6	873	4	US-09-470-946-6
6	629.5	24.6	873	4	US-09-438-906-2
7	629.5	24.6	873	4	US-09-438-906-4
8	629.5	24.6	925	2	US-08-392-946-1
9	629.5	24.6	925	2	US-08-504-165-1
10	629.5	24.6	925	5	PCT-US94-14893-1
11	534.5	20.9	438	3	US-09-187-331-2
12	534.5	20.9	438	4	US-09-470-946-2
13	465.5	18.2	829	1	US-08-346-455B-34
14	465.5	18.2	829	3	US-08-977-221-34
15	465.5	18.2	829	4	US-09-483-831B-34
16	465.5	18.2	829	5	PCT-US95-06613-34
17	465.5	18.2	915	3	US-08-346-455B-69
18	465.5	18.2	915	3	US-08-977-221-69
19	465.5	18.2	915	4	US-09-483-831B-69
20	465.5	18.2	915	5	PCT-US95-06613-69
21	429.5	16.8	861	1	US-08-346-455B-67
22	429.5	16.8	861	3	US-08-977-221-67
23	429.5	16.8	861	4	US-09-483-831B-67
24	429.5	16.8	861	5	PCT-US95-06613-67
25	375	14.6	788	1	US-08-346-455B-36
26	375	14.6	788	3	US-08-977-221-36
27	375	14.6	788	4	US-09-483-831B-36

28	375	14.6	788	5	PCT-US95-06613-36	Sequence 36, Appl
29	371	14.5	979	1	US-08-346-455B-38	Sequence 38, Appl
30	371	14.5	979	3	US-08-977-221-38	Sequence 38, Appl
31	371	14.5	979	4	US-09-483-831B-70	Sequence 38, Appl
32	371	14.5	979	5	PCT-US95-06613-38	Sequence 38, Appl
33	197	7.7	151	4	US-09-621-976-3891	Sequence 3891, Ap
34	139.5	5.4	108	4	US-09-621-976-7142	Sequence 7142, Ap
35	111.5	4.4	589	4	US-09-543-681A-4194	Sequence 4194, Ap
36	109	4.3	709	4	US-09-668-673B-3	Sequence 3, Appl
37	105.5	4.1	819	4	US-09-468-656A-10	Sequence 10, Appl
38	101.5	4.0	972	3	US-08-335-844A-23	Sequence 23, Appl
39	101.5	4.0	972	4	US-09-129-366-23	Sequence 23, Appl
40	101	3.9	627	4	US-09-328-352-7547	Sequence 7547, Ap
41	98.5	3.8	763	3	US-08-361-083-66	Sequence 66, Appl
42	98.5	3.8	763	4	US-09-536-784-66	Sequence 66, Appl
43	98.5	3.8	838	4	US-09-468-656A-4	Sequence 4, Appl
44	98	3.8	1541	3	US-08-296-791-3	Sequence 3, Appl
45	98	3.8	1541	4	US-09-839-996-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-014-969-19
Sequence 19, Application US/09014969
Patent No. 5965397

GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Mebery, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Yakkil
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 488-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-014-969-19

Query Match 70.7%; Score 1810; DB 2; Length 355;
Best Local Similarity 99.4%; Pred. No. 2.6e-173;
Matches 335; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY      1 MTSKFLVSLFLLAALSLSTFSLQDQOKVLVSPFGFPMYLYKVPFPHYIMKXGVH 60
DB      1 MTSKFLVSLFLLAALSLSTFSLQDQOKVLVSPFGFPMYLYKVPFPHYIMKXGVH 60
QY      61 VKQVNVFITYKYNHYTLVTLGFAENHGIYANDMDPIRNKSFSLDHNIYDSKFEWA 120
DB      61 VKQVNVFITYKYNHYTLVTLGFAENHGIYANDMDPIRNKSFSLDHNIYDSKFEWA 120
QY      121 TPWITNORAGHTSGAAMPGTDVYKHKRPFTHYMESVSPEDRAKIVEMFTSKPEI 180
DB      121 TPWITNORAGHTSGAAMPGTDVYKHKRPFTHYMESVSPEDRAKIVEMFTSKPEI 180
QY      181 NGLLIYMEDPDMGHHLGPDSPIMGVPSIDIKKGLYLQMLKAKLMTNLIIITSDHG 240
DB      181 NGLLIYMEDPDMGHHLGPDSPIMGVPSIDIKKGLYLQMLKAKLMTNLIIITSDHG 240
QY      241 MPOGSEERLISDQYLDKDHVTLIDOSPVAIIPKRGKRDVYALTHAHPNLYKKED 300
DB      241 MPOGSEERLISDQYLDKDHVTLIDOSPVAIIPKRGKRDVYALTHAHPNLYKKED 300
QY      301 VPERMHHYKNSRIQPIIAVADSGWHIIQNSDDFLVG 337
DB      301 VPERMHHYKNSRIQPIIAVADSGWHIIQNSDDFLVG 337

```

RESULT 2

```

US-09-800-729-83
; Sequence 83, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-83

```

Query Match

```

40.4%; Score 1034; DB 4; Length 453;
Best Local Similarity 45.9%; Pred. No. 3,7e-95;
Matches 208; Conservative 73; Mismatches 132; Indels 40; Gaps 10;

```

```

QY      6 IIVSFLIALLSLSTFSLQDQOKVLVSPFGFPMYLYKVPFPHYIMKXGVH 65
DB      7 LIPSGILTCGSSSHL--PSKLLVSPFGFADLTQNYFPLQNFIKESGVLEHAK 63
QY      66 NVFITYKYNHYTLVTLGFAENHGIYANDMDPIRNKSFSLDHNIYDSKFEWA 122
DB      64 NVFITYKYNHYTLVTLGFAENHGIYANDMDPIRNKSFSLDHNIYDSKFEWA 118
QY      123 IWTNQG-RAGHTSGAAMPGTDVYKHKRPFTHYMESVSPEDRAKIVEMFTSKPEI 180
DB      119 IWTNQG-RAGHTSGAAMPGTDVYKHKRPFTHYMESVSPEDRAKIVEMFTSKPEI 178
QY      181 NGLLIYMEDPDMGHHLGPDSPIMGVPSIDIKKGLYLQMLKAKLMTNLIIITSDHG 239
DB      179 TPWITNORAGHTSGAAMPGTDVYKHKRPFTHYMESVSPEDRAKIVEMFTSKPEI 238
QY      240 GMTQSGERLISDQYLDKDHVTLIDOSPVAIIPKRGKRDVYALTHAHPNLYKKED 299
DB      239 GMTQSGERLISDQYLDKDHVTLIDOSPVAIIPKRGKRDVYALTHAHPNLYKKED 297
QY      300 VPERMHHYKNSRIQPIIAVADSGWHIIQNSDDFLVG 359

```

Hed A.

```

DB      298 DIPARFHYQANDRIQPIIAVADSGWHIIQNSDDFLVG 356
QY      360 RKNFSKEMNSTDLVPLCHLNTTAMPHNGSSFWNVQDLNLSAMERVVPTOSTILLPGS 419
DB      357 HKQYKHSITNSVDVYPMCMCHILGLKPHENNGTGTAKKCL-----VDQWCIINLEA 407
QY      420 VKPAEYDQSGSYPIFGVSLGSIYIVFVFI 452
DB      408 -----IGIVGALVLTLLTCL 425

```

RESULT 3

```

US-09-621-976-3913
; Sequence 3913, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 3913
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FRAGMENT:
; NAME/KEY: SIGNAL
; LOCATION: -22..-1
US-09-621-976-3913

```

Query Match

```

28.1%; Score 719; DB 4; Length 136;
Best Local Similarity 99.3%; Pred. No. 2.4e-64;
Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 MTSKFLVSLFLLAALSLSTFSLQDQOKVLVSPFGFPMYLYKVPFPHYIMKXGVH 60
DB      1 MTSKFLVSLFLLAALSLSTFSLQDQOKVLVSPFGFPMYLYKVPFPHYIMKXGVH 60
QY      61 VKQVNVFITYKYNHYTLVTLGFAENHGIYANDMDPIRNKSFSLDHNIYDSKFEWA 120
DB      61 VKQVNVFITYKYNHYTLVTLGFAENHGIYANDMDPIRNKSFSLDHNIYDSKFEWA 120
QY      121 TPWITNORAGHTSGA 136
DB      121 TPWITNORAGHTSGA 136

```

RESULT 4

```

US-09-187-331-6
; Sequence 6, Application US/09187331
; Patent No. 6043056
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Gorgone, Gina A.
; APPLICANT: Baughn, Marian R.
; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
; FILE REFERENCE: PF-0631 US
; CURRENT APPLICATION NUMBER: US/09/187,331
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -

```